A Heuristic Approach to Optimise Public Transport Priority in an Urban Network

Mahmoud Mesbah, Majid Sarvi
Institute of Transport Studies, Department of Civil Engineering, Monash University, Victoria, Australia 3800

Traffic congestion is a challenge facing many urban transport networks. While construction of new roads is not a viable option in many cities, improvement of public transport (PT) facilities can offer a solution. Introduction of exclusive lanes for PT vehicles is a green and low cost improvement which is focused in this paper. Many approaches in the literature devised an exclusive lane based on a local perspective and some other studies evaluated priority lanes on urban freeways. It is of great importance to develop methods for introduction of PT exclusive lanes from a network perspective and on arterial roads. In this study, a method is proposed to locate PT exclusive lanes in an urban transport network. Considering all the stakeholders of the transport network, a comprehensive transport planning model is used to evaluate a given set of exclusive lanes. In the transport planning model, a traffic assignment model and a transit assignment model is included to consider the primary effect of the prioritization, as well as a modal split model to capture the secondary effects of mode shift. A Genetic Algorithm approach is applied to find the optimal solution. The procedure is implemented by a powerful interface with a Visum and computational issues are addressed for a real world application.

INTRODUCTION

Bus is the main mode of Public Transport (PT) in many cities (Hensher, 1999). Flexibility of using the road network, low infrastructure cost, and high access are among the reasons which make bus mode attractive. However, low operational speed is the major drawback of this mode as it usually runs in a mixed traffic condition. A solution to this low speed issue is to introduce a bus exclusive lane (Luk, 1992). This paper focuses on the allocation of roadspace to bus and private car modes in order to manage the existing transport network. Since buses move higher number of passengers than private cars, the efficiency of a road can be increased if the road-space is shared appropriately (Currie et al., 2007). This is the concept of exclusive bus lane which can be used as a powerful mean in network management (Luk, 1992). While a wide range of studies have recommended criteria for allocation of a lane to bus vehicles, they can be divided to studies with local and network perspectives.

One part of studies have focused on a link or a corridor. Black et al. (1992) presented a model to evaluate several alternative road space allocations for a corridor. The total travel cost of users in the corridor was considered as the performance measure. In another attempt, Jepson and Ferreira (2000) assessed different road space priority treatments such as a bus lane and set-backs based on delays in two consecutive links. Using the concept of intermittent bus lanes (Viegas and Lu, 2004), Eichler and Daganzo (2006) suggested an analysis method which is based on kinematic wave theory. This method can be applied to a long arterial.

The above researchers have focused on examining bus lane problems on an individual link level basis. Only a few researchers have considered the problem from a network wide, multi link viewpoint. Waterson et al. (2003) represented a macro-simulation approach which evaluates a given priority scenario in the network of Southampton city. This approach considered rerouting, retiming, modal change, and trip suppression. A similar evaluation approach is carried out using a micro-simulation by Liu et al. (2006). Stirzaker and Dia (2007) applied another micro-simulation approach to evaluate a major bus lane project in Brisbane.

Although the reviewed studies have different focuses in terms of the spread of the proposed exclusive lanes, all researches evaluate some given transit priority alternatives (TPAs). Despite the great level of details in some studies, the evaluation just reveals weather or not a TPA (i.e. a set of bus exclusive lanes) should be implemented. It does not mean that the given TPA is the best possible or optimum TPA for the network. Therefore, it is necessary to propose an optimization method to find the best alternative of bus lanes.

This paper outlines a methodology to find the optimal TPA. The optimal TPA determines the links in the transport network on which an exclusive bus lane should be introduced. Furthermore it is aimed at presenting a methodology that can be applied to medium and large size networks. In the next section, the optimization method is formulated as a bi-level programming problem and each level is explained separately. Then, in section 3 a solution algorithms is presented based on Genetic Algorithm. In the last section the optimization problem is solved for a medium size network and the results are presented.

BI-LEVEL OPTIMIZATION

There are two levels of stakeholders who determine the performance of a transit priority scheme. At the upper level, the transport authority would propose a TPA which is a set of links on which priority is provided. Given this TPA, at the lower level, system users would choose a strategy to maximize their own benefit under the prevailing conditions. This problem can be modelled as a Stackelberg competition where the transport authority is the leader and system users are the followers (Yang and Bell, 1998). In equilibrium conditions, the optimal TPA is chosen. The Stackelberg model can be modelled as a bi-level optimization.

Transport authority’s point of view is considered at the upper level. Therefore a system optimal is formulated in this paper for the upper level. The transport authority takes into account the total travel time of car and transit users as well as other performance measures of the system such as travel cost and emissions. There can also be a series of practical constraints for a priority scheme which is formulated in the constraints of the upper level. In the next subsection, a comprehensive objective function and associated constraints are defined. The output of the upper level is the set of decision variables which define the location of the exclusive lanes.

User behaviour at the lower level is modelled by applying a traditional four step modelling. In this study, it is assumed that the total travel demand in the network is not changed by introduction of a TPA. It is also assumed that two modes of private car and bus use the network. Thus, the total demand should be split to these two modes. In the last step of planning, car and bus demand should be assigned to the network links. At the lower level for private cars and buses, a traffic assignment model and a transit assignment model are used, respectively. It is important to note that the TPA is determined at the upper level while it is in the lower level where the objective function can be calculated. The formulation of the lower level is discussed in the subsequent sections.

UPPER LEVEL FORMULATION

A system optimal from the transport authority's perspective is formulated at the upper level. This formulation was first proposed by Mesbah et al. (2009) to consider benefits of car and bus users, as well as impact of a TPA on environmental measures of the network. The upper level can be proposed as follows.

$$MinZ = \alpha \sum_{a \in A} x_a^c t_a^c(x) + \beta(\sum_{a \in B} x_a^b t_a^b(x) + \sum_{i \in I} w_i) + \gamma \sum_{a \in A} \frac{x_a^c}{Occ^c} l_a Imp^c + \eta \sum_{a \in B} f_a l_a Imp^b \qquad (1)$$

Subject to :

$$\sum_{a \in A_2^{'}} Exc_a \phi_a \leq Bdg \qquad (2)$$

$$\phi_a = 0 \text{ or } 1 \quad \forall a \in A_2 \qquad (3)$$

where:

A : Set of all links in the network, $A = A_1 \cup A_2 \cup A_2^{'}$

A_1 : Set of links in the network where provision of priority is impossible,

A_2 : Set of links with priority lane (with exclusive lane),

$A_2^{'}$: Set of conjugate links with mixed traffic (no exclusive lane),

B : Set of links with a bus line on them, walking links, and transfer links,

L : Set of bus lines,

I : Set of bus stops,

f_a : Sum of frequency of service for all bus lines on link (a),

f_p : Frequency of service for bus line p,

l_a : Length of link (a),

$t_a^{c,b}(x)$: Travel time on link (a) by car (c) or bus (b) which is a function of flow,

w_i : Total waiting time for users at node (i),

$x_a^{c,b}$: Passenger flow on link (a) by mode car (c) or bus (b),

Bdg : Available budget,

Exc_a: The cost of implementation of an exclusive lane on link (a),

$Imp^{c,b}$: Aggregate weight of operation costs of a car (c) or bus (b) to the community per km including: emissions, noise, accident, and reliability impacts.

Occ^c : Average occupancy rate for the car mode,

$\alpha, \beta, \gamma, \eta$: weighting factors to convert the units and adjust the relative importance of each impact in the objective function, $\alpha, \beta, \gamma, \eta \geq 0$,

ϕ_a : Equals to 1 if there is an exclusive lane on link (a), and 0 otherwise, Note that $f_a = \sum_{p \in L} f_p \xi_{p,a}$ where $\xi_{p,a}$ is the bus line-link incident matrix and $t_a^b(x)$ is the in-vehicle travel time.

The first term of the objective function is the total travel time by car; while the second term represents the total travel time by bus including access time, waiting time and transfer time. The next two terms are corresponded to the cost of total travel distance by car and bus. Coefficients α,β,γ,η can reflect different policies in the relative importance of each term. They also convert the units. As the Equation **(1)** shows, the objective function is formed from transport authority's perspective. The budget constraint is demonstrated in Equation **(2)**.

There are two types of links in the network. The first type is the links on which no lane can be dedicated to buses. This type includes collector links and links with special consideration. The second type is the links that potentially can have an exclusive lane. In the network model, instead of each link in this type, two links should be defined: one with and one without an exclusive bus lane (sets A_2 and $A_2^{'}$). Decision variables determine which links would actually be in the real network. Noticeably, only one of the links can be selected. Based on the set of decision variables in the upper level, flow and travel time is computed at the lower level.

LOWER LEVEL FORMULATION

When a TPA is determined, it is users turn to decide on how they would utilize the provided supply. In other words models at the lower level estimate users response to a given TPA. These models in the bi-level structure function as constraints to the optimization programming presented in the upper level. As a result of these models flow and travel time is obtained.

Assuming a constant travel demand, as discussed before, there are 3 models involved in the transport modelling:

1. Modal split
2. Traffic assignment
3. Transit assignment

The modal split model predicts the share of car and bus in travel demand. For this purpose, a Logit model (Ortúzar and Willumsen, 2001) is applied. The model calculates a utility function for each mode of travel from its attributes. Then, the probability of travelling by a mode is found depending on the utility value. Since two modes of travel are available, two utility functions should be developed. Priority provision can shift the travel demand to use bus. A change in the decision variables changes the attributes of travel by each mode which in turn can influence the mode share.

Traffic assignment is the second model at the lower level. Traffic assignment is carried out using a static User Equilibrium (UE) model which is a conventional model for strategic planning (Sheffi, 1984). This model finds car flow and travel time in the network by an optimization approach. The effect of the decision variables in the flow and travel time can not explicitly be expressed; this is one of the reasons that a bi-level approach is proposed. The decision variables of the upper level optimization would appear at the constraints of the UE formulation as follows:

$$MinY = \sum_{a \in A} \int_0^{x_a^c} t_a^c(x)dx \tag{4}$$

Subject to:

$$\sum_k f_k^{rs} = q_{rs}^c \qquad \forall r,s \tag{5}$$

$$f_k^{rs} \geq 0 \qquad \forall k,r,s \tag{6}$$

$$x_a^c = \sum_{rs} \sum_k f_k^{rs} \delta_{k,a}^{rs} \qquad \forall (i,j) \in A \tag{7}$$

$$x_a^c \leq M\phi_a \qquad \forall (i,j) \in A_2 \tag{8}$$

$$x_a^c \leq M(1-\phi_a) \qquad \forall (i,j) \in A_2^{'} \tag{9}$$

where f_k^{rs} is the flow on path k connecting origin node r to destination node s, q_{rs} is the trip rate between r and s, and x_a^c is related to f_k^{rs} by the incident matrix $\delta_{k,a}^{rs}$ where δ is 1 if link (a) is on path k for any OD pair rs and zero otherwise; M is a big enough constant.

In the above constraints, the first two are conservation of flow and non negativity constraints. The third constraint defines the relation of paths to links. The next two constraints (Equations **(8)** and **(9)**) prevent traffic flow on the links which are not being used. The decision variables on the right hand side of the Equations **(8)** and **(9)** bind the two levels together. Instead of each candidate link, two links are defined (see section 2). The binding constraints ensure that only one of these coupled links would have a positive flow.

Transit assignment is the third model to be used which assigns the bus demand to the transport network. Transit assignment is the second reason for which a bi-level approach is proposed. All of the models proposed in the literature for transit assignment can be applied in this framework. Nevertheless, some binding constraints similar to Equations **(8)** and **(9)** should be added to their formulation. In this paper, a model based on Spiess and Florian (1989) is adapted.

$$MinW = \sum_{a \in B} x_a^b t_a^b + \sum_{i \in I} w_i \tag{10}$$

Subject to:

$$\sum_{a \in B_i^+} x_a^b - \sum_{a \in B_i^-} x_a^b = q_i^b \qquad \forall i \in I \tag{11}$$

$$x_a^b \leq f_a w_i \quad \forall a \in B_i^+,\ \forall i \in I \tag{12}$$

$$x_a^b \leq M\phi_a \qquad \forall a \in A_2 \tag{13}$$

$$x_a^b \leq M(1-\phi_a) \qquad \forall a \in A_2^{'} \tag{14}$$

$$x_a^b \geq 0 \qquad \forall a \in B \tag{15}$$

where $B_i^{+/-}$ is the set of outgoing/incoming links (incoming with negative sign) from/to node i, q_i^b is the demand at node i. It is assumed that the stops are located on the nodes. The first constraint is conservation of flow and the second constraint divides flow proportional to the frequency of links. Equations **(13)** and **(14)** are the abovementioned binding constraints. The last constraint ensures non-negativity of flow.

SOLUTION ALGORITHM

Bi-level structure even with linear objective function and constrains at both levels is a NP-hard problem and difficult to solve. In this study a heuristic approach based on Genetic Algorithm (GA) proposed in which the new answers are produced by combining two predecessor answers (Russell and Norvig, 2003). Inspired from evolutionary theory in the nature, GA starts with a feasible set of answers called population (see Figure 1). Each individual answer in the population (called a chromosome) is assigned a survival probability based on the value of the objective function. Then, the algorithm *selects* individual chromosomes based on this probability to breed the next generation of the population. GA uses *cross over* and *mutation* operators to breed the next generation which replaces the predecessor generation. The algorithm is repeated with the new generation until a convergence criterion is satisfied. A number of studies applied GA to transit networks. Two recent examples are a transit network design problem considering variable demand (Fan and Machemehl, 2006) and minimization of transfer time by shifting time tables (Cevallos and Zhao, 2006).

In this study, GA is applied to optimize transit priority. To adapt GA to the concept of this study, a gene is defined to represent the binary variable ϕ and a chromosome is the vector of genes (Φ) which is a TPA. A chromosome (or TPA) contains a feasible combination of links on which an exclusive lane may be introduced (set A_2). Therefore, the length of the chromosome is equal to the size of A_2. The algorithm starts with a feasible initial population. The chromosomes of the initial population are produced randomly. To ensure the feasibility, according to constraint represented in equation **(2)**, the cost of each chromosome is calculated. If the cost exceeds the budget, one of the genes with value of 1 is changed to 0 and this continues until the cost becomes less than the budget.

Once a feasible chromosome population is produced, the upper level objective function for all chromosomes should be determined. Each chromosome identifies the leader’s decision vector for the network. It is users’ turn at the lower level to choose their route. Thus, for each chromosome, the lower level models are carried out as depicted in Figure 1. Using the flow and travel time at the lower level, the objective function for the chromosome is determined. The lower level calculations are repeated for all chromosomes in the population (Figure 1).

The chromosomes with higher value of the objective function are assigned a higher survival probability. Then, GA operators of selection, cross over, and mutation are employed to produce the next generation (set of TPAs). Similar to the process in the initial population, the process ensures the feasibility of the new generation. The new generation is replaced the previous one and the calculations are repeated. It should be noted that to increase the convergence rate of the algorithm (it is recommended that) the best chromosome of the previous population is kept. The algorithm stops when either the number of iterations reaches the maximum number of iterations or the best answer does not improve in a certain number of iterations. This cycle is also demonstrated in Figure 1.

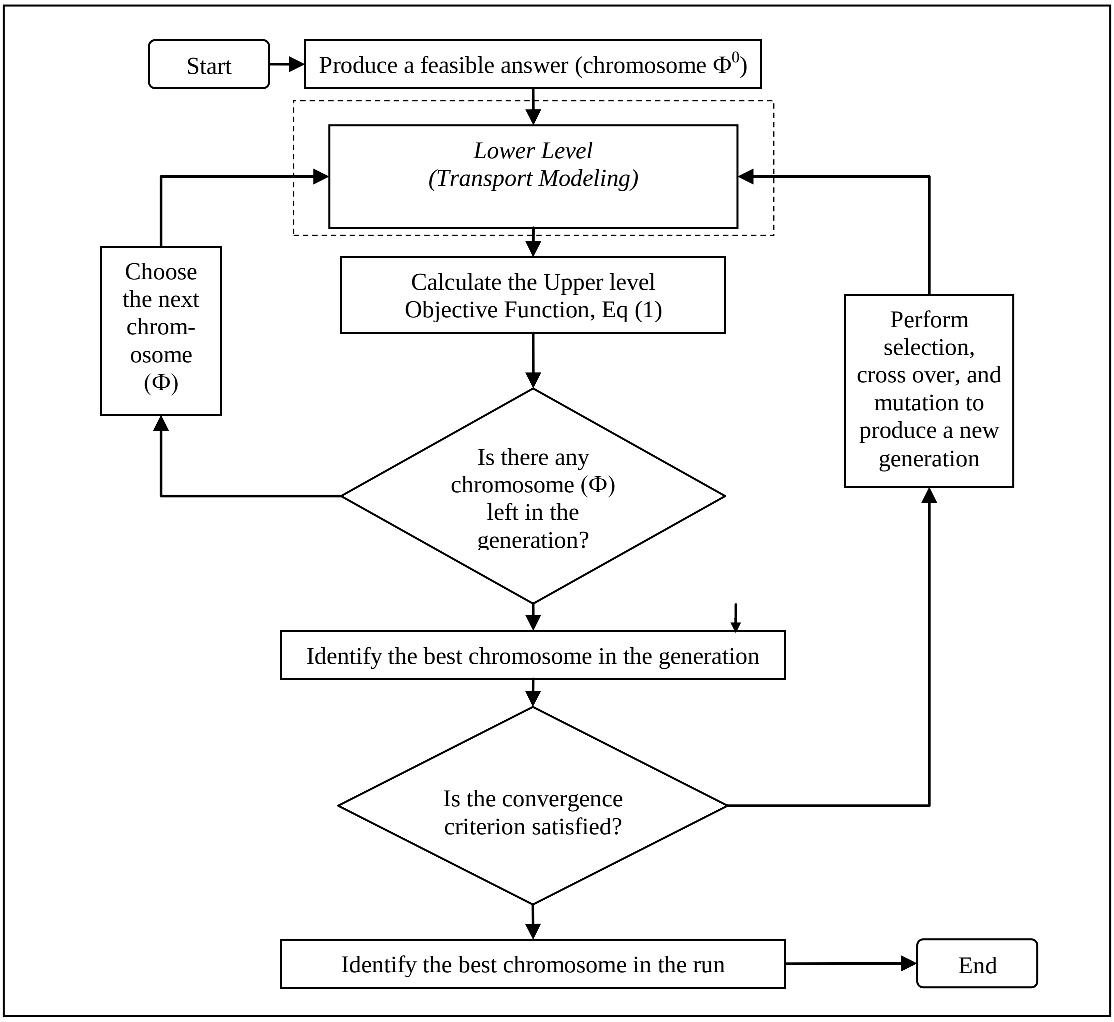

Figure 1 GA Solution Flowchart

EXAMPLE

In this section, the proposed method is applied to an example network. Figure 2 shows the layout of the network. This grid network consists of 86 nodes and 306 links. All the circumferential nodes together with Centroid 22, 26, 43, 45, 62, and 66 are origin and destination nodes. A flat demand of 30 Person/hr is traveling from all origins to all destinations. The total demand for all the 36 origin destination is 37800 Person/hr. There are 10 bus lines covering transit demand in the network as shown in Figure 2. The frequency of service for all the bus lines is 10 min. The models and parameters used in this example are extracted form those calibrated for Melbourne Integrated Transport Model (MITM) (Department of Infrastructure, 2004).

Vertical and horizontal links are 400m long with two lanes in each direction and the speed limit of 36 km/hr. It is assumed that if an exclusive lane is introduced on a link, it would also be introduced on the other link with opposite direction. There are a total number of 120 links (60 links two directional) in the network of Figure 2 on which an exclusive lane can be introduced. These links are highlighted in Figure 2. The following Akcelik cost functions (Ortúzar and Willumsen, 2001) are assumed for links with exclusive lane (Equation (16)) and without exclusive lane (Equation (17)):

$$t^c_{1,a} = t_{0,a} + \frac{3600a}{4}\left[\left(\frac{x^c_a}{Cap^c_{1,a}} - 1\right) + \sqrt{\left(\frac{x^c_a}{Cap^c_{1,a}} - 1\right)^2 + \frac{8b}{ad}\left(\frac{x^c_a}{Cap^c_{1,a}}\right)}\right], t^t_{1,b} = t_{0,a} \qquad 16)$$

$$t_{0,a}^{c} = t_{0,a}^{b} = t_{0,a} + \frac{3600a}{4}[(\frac{x_a^c + x_a^t}{Cap_{1,a}^c} - 1) + \sqrt{(\frac{x_a^c + x_a^t}{Cap_{1,a}^c} - 1)^2 + \frac{8b}{ad}(\frac{x_a^c + x_a^t}{Cap_{1,a}^c})}] \quad 17)$$

where t_0 determines travel time with free flow speed, a is length of observation period, b is a constant, d is lane capacity, and other terms are as defined earlier. It is assumed that each link has 2 lanes and:

$$a = 1hr, b = 1.4, d = 800veh/hr$$

$$Cap_{0,a}^{c} = 1800veh/hr$$

$$Cap_{1,a}^{c} = 900veh/hr$$

Mode share is determined using a Logit model (Equation (4)-(5)). In Equation (5), the average travel time (X_1) and distance (X_2) between origin destination node pairs are considered for mode attribute. It is also assumed that:

$$U^c = 2 + 0.2731 * X_1^c + 0.2235 * X_2^c$$

$$U^b = 3.9 + 0.1300 * X_1^b$$

Once the demand matrices are determined, car demand is assigned using UE and bus demand is assigned using frequency based assignment. It is assumed that bus frequencies are fixed in this example. The feedback process from assignment to modal split is performed to adjust the assumed attributes in Modal split. The convergence criterion of the feedback process is set on the difference of the travel time on a link from its travel time in the last iteration. The lower level transport model is implemented using Visum modeling package (PTV AG, 2007).

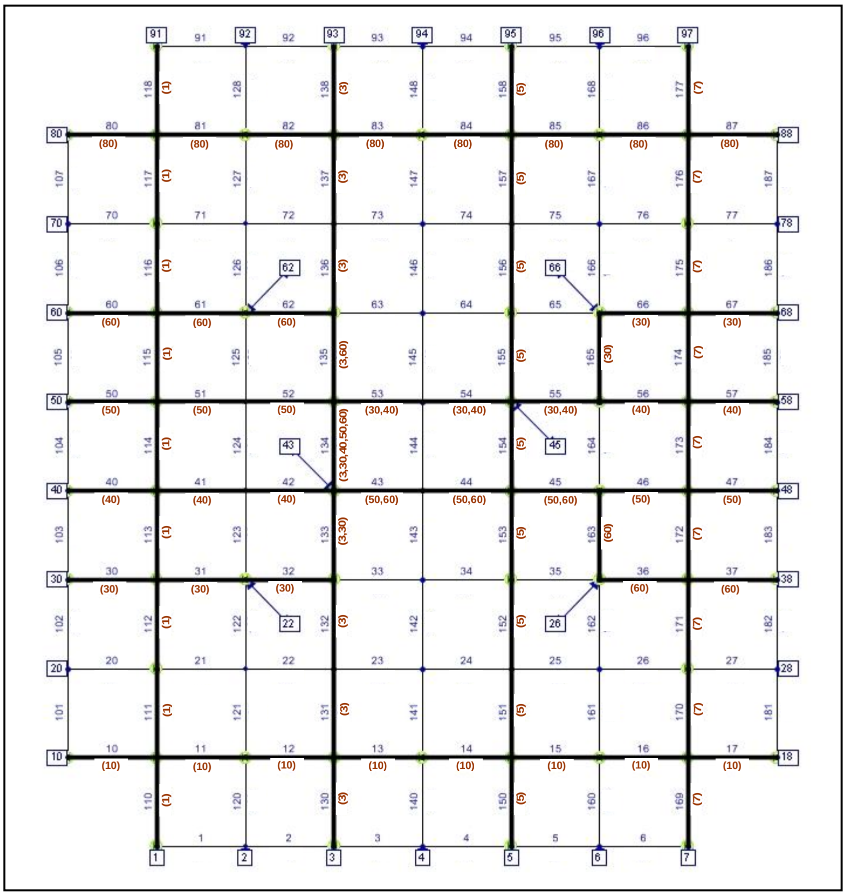

Figure 2 Example network with origin destination nodes in box and bus lines in parenthesis.

In this example weighting factors of the upper level objective function are assumed equal. These factors may vary depending on the relative importance of the factors in viewpoint of transport authorities. The upper level objective function includes total travel time (veh.sec) and total vehicle distance (veh.km). The absolute value of the objective function therefore can be very large. In order to avoid numerical problems, the improvement of each term compared to a base case is considered instead of the absolute value of the term in the objective function. The base case is assumed to be the case where no link is provided with an exclusive lane $(\Phi = 0)$. Regarding constraints, it is assumed that budget allows for all candidate links for the provision of bus priority.

A common stopping criterion for GA is the number of generations. If the objective function does not improve for a considerable number of generations, the calculations are terminated. In this example, the number of generations is increased to 2000 to investigate a proper stopping criterion. Figure 3 demonstrates the value of the objective function for two independent runs of GA. As this figure shows the objective function did not improve after 800 generations which can be introduced as the stopping criterion for this example.

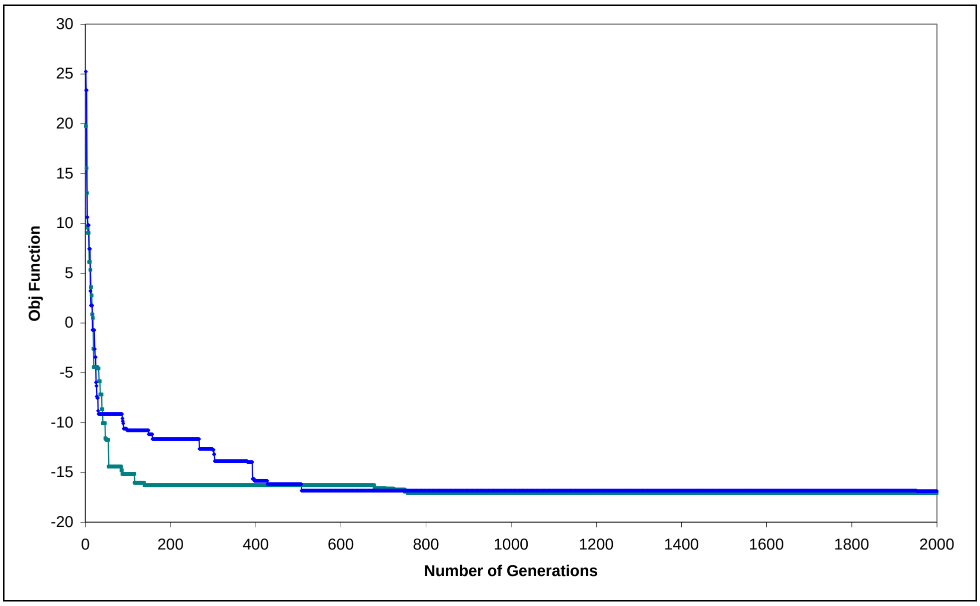

Figure 3 Effect of Number of Generations on the Value of the Objective Function

Application of the proposed method to the network of Figure 2 resulted in introduction of an exclusive bus lane on the following 22 links.

31,32,36,41,43,44,45,53,54,55,56,61,62,66,131,132,133,134,135,136,137,154

This answer is anticipated since the it includes all links on which two or more bus lines were travelling (134,43,44,45,53,54,55,133,135). It also includes links 131 to 136 which is the busiest north-south bus corridor and excludes outer links with low bus patronage such as 11 to 16 and 111 to 117.

CONCLUSION

A heuristic approach to optimise public transport priority is presented in this paper. It was stated that all the previous approaches consider a limited number of alternatives for transit exclusive lanes while all the feasible combinations are taken into account in the proposed approach. The problem is elaborated in a framework of bi-level programming formulation: The upper level being system optimal from the transport authority's perspective and the lower level a modified four step modelling to predict user's behaviour. An efficient solution algorithm based on Genetic Algorithm is suggested to solve the bi-level optimization. The method is applied to a medium size example network and the results are presented. The proposed method should also be tested in a real scale network with addition of practical constraints at the upper level.

REFERENCES

BLACK, J. A., LIM, P. N. & KIM, G. H. (1992) A traffic model for the optimal allocation of arterial road space: a case study of Seoul's first experimental bus lane. *Transportation Planning and Technology,* 16**,** 195-207.

CEVALLOS, F. & ZHAO, F. (2006) Minimizing Transfer Times in Public Transit Network with Genetic Algorithm. *Transportation Research Record***,** pp 74-79.

CURRIE, G., SARVI, M. & YOUNG, B. (2007) A new approach to evaluating on-road public transport priority projects: Balancing the demand for limited road-space. *Transportation,* 34**,** 413-428.

DEPARTMENT OF INFRASTRUCTURE (2004) Melbourne Multi-modal Integrated Transport Model (MITM), User Guide.

EICHLER, M. & DAGANZO, C. F. (2006) Bus lanes with intermittent priority: Strategy formulae and an evaluation. *Transportation Research Part B: Methodological,* 40**,** 731-744.

FAN, W. & MACHEMEHL, R. B. (2006) Optimal Transit Route Network Design Problem with Variable Transit Demand: Genetic Algorithm Approach. *Journal of Transportation Engineering,* 132**,** pp 40-51.

HENSHER, D. A. (1999) bus-based transitway or light rail? Continuing the saga on choice versus blind commitment. *Road and Transport Research,* 8.

JEPSON, D. & FERREIRA, L. (2000) Assessing travel time impacts of measures to enhance bus operations. Part 2: Study methodology and main findings. *Road and Transport Research,* 9**,** 4-19.

LIU, R., VAN VLIET, D. & WATLING, D. (2006) Microsimulation models incorporating both demand and supply dynamics. *Transportation Research Part A: Policy and Practice,* 40**,** 125-150.

LUK, J. (1992) Models for travel demand management - a review. Australian Road Research Board Ltd (ARRB).

MESBAH, M., SARVI, M., OUVEYSI, I. & CURRIE, G. (2009) Optimization of Transit Priority in the Transportation Network Using a Decomposition Methodology. *the 18th International Symposium on Traffic and Transport Theory (ISTTT).* Hong Kong.

ORTÚZAR, J. D. D. & WILLUMSEN, L. G. (2001) *Modelling transport,* Chichester New York, J. Wiley.

PTV AG (2007) VISUM 9.5 User Manual. 9.5 ed. Karlsruhe, Germany.

RUSSELL, S. J. & NORVIG, P. (2003) *Artificial intelligence : a modern approach,* Upper Saddle River, N.J., Prentice Hall.

SHEFFI, Y. (1984) *Urban transportation networks: equilibrium analysis with mathematical programming methods,* Englewood Cliffs, N.J., Prentice-Hall.

SPIESS, H. & FLORIAN, M. (1989) Optimal strategies. A new assignment model for transit networks. *Transportation Research, Part B: Methodological*, 23B**,** 83-102.

STIRZAKER, C. & DIA, H. (2007) Evaluation of transportation infrastructure management strategies using microscopic traffic simulation. *Journal of Infrastructure Systems,* 13**,** 168-174.

VIEGAS, J. & LU, B. (2004) The Intermittent Bus Lane signals setting within an area. *Transportation Research Part C: Emerging Technologies,* 12**,** 453-469.

WATERSON, B. J., RAJBHANDARI, B. & HOUNSELL, N. B. (2003) Simulating the impacts of strong bus priority measures. *Journal of Transportation Engineering,* 129**,** 642-647.

YANG, H. & BELL, M. G. H. (1998) Models and algorithms for road network design: a review and some new developments. *Transport Reviews* 18**,** 257-278.